Modified Gaussian Approaches for Peak Deconvolution in Chiral Liquid Chromatography: Analytical Method Development

Dr. Claire Dubois[1]*, Dr. Antoine Lefèvre[1], Dr. Sophie Martin[1]

[1]Sorbonne University, Department of Analytical Chemistry and Chromatographic Science, Paris, France

Abstract

The description of the profiles of chromatographic peaks has been studied extensively, with a large number of proposed mathematical functions. Among them, the accuracy achieved with modified Gaussian models that describe the deviation of an ideal Gaussian peak as a change in the peak variance or standard deviation over time, has been highlighted. These models are, in fact, a family of functions of different complexity with great flexibility to adjust chromatographic peaks over a wide range of asymmetries and shapes. However, an uncontrolled behaviour of the signal may occur outside the region being fitted, forcing the use of different strategies to overcome this problem. In this work, the performance of the LMG (Linear Modified Gaussian), PVMG (Parabolic Variance Modified Gaussian), and PLMG (Parabolic-Lorentzian Modified Gaussian) models is compared with variants obtained by combination of the modified Gaussian models with an equation that adds an exponential tail and with other functions that limit the growth of the independent variable. The behaviour of the approaches is checked through the simultaneous fitting of enantiomeric peaks showing a wide range of characteristics, obtained in the separation of drugs with chiral activity by liquid chromatography using enantioselective columns. The study is also carried out with the purpose of performing the deconvolution of the peaks of the enantiomers, when these are not completely resolved, in order to evaluate the enantiomeric fraction.

Keywords: Liquid chromatography; Enantiomeric analysis; Peak fitting; Modified Gaussian models; Peak deconvolution

*Corresponding author. E-mail: celia.garcia@uv.es (M.C. García-Alvarez-Coque).

1. Introduction

In order to forecast the experimental conditions that best resolve the components of a sample by liquid chromatography, in the shortest analysis time, it is convenient to make realistic predictions of the expected chromatograms. A reliable methodology to make such predictions is based on the use of models that describe the retention and shape of chromatographic peaks exhibiting different widths and asymmetries [1–5]. The parameters of the peak function used to fit the peaks should offer information about their position (t_R), height (H_0), and left (A) and right (B) half-widths [6,7]. Also, models able to predict the peak half-widths as a function of the experimental conditions are very useful for the simulation of chromatograms [8,9].

A large number of mathematical functions have been proposed to describe chromatographic peaks with different success [10–19]. Of special interest is the compilation of equations by Di Marco and Bombi [16]. The exponentially modified Gaussian function has been especially popular and served Foley and Dorsey as a basis to obtain different equations to calculate parameters that characterise asymmetric chromatographic peaks, depending on the values of the left and right half-widths, A and B, and the retention times [20,21]. Our research group has proposed modified Gaussian models [6,13,22], which have demonstrated very good performance to describe peaks with diverse asymmetry, where the model parameters can be expressed as a function of the half-widths [7].

The simplest model (the linearly modified Gaussian, LMG model) uses a linear function to describe the peak standard deviation [13]:

$$h(t) = H_0 \exp\left[-\frac{1}{2}\left(\frac{t_c}{\sigma_0 + m\,t_c}\right)^2\right] \quad (1)$$

being

$$\sigma_0 = 0.932\,\frac{A\,B}{A+B} \quad (2)$$

$$m = 0.466 \ \frac{B - A}{A + B} \tag{3}$$

where $t_c = t - t_R$, and the half-widths are measured at 10% peak height. According to Foley and Dorsey [11,12], measurements at this height are most convenient, since peak asymmetry is more evident and the baseline noise is not disturbing. In the LMG model, σ_0 describes the standard deviation of a Gaussian that behaves as the asymmetric peak at the maximum [7]. This model is ideal for predicting peaks with moderate asymmetries.

For more accurate fitting of chromatographic peaks, a more complex model, such as the modified Gaussian with a parabolic function (PVMG model) [6] is needed:

$$h(t) = H_0 \ \exp\left[-\frac{1}{2} \ \frac{t_c^2}{\sigma_0^2 + b\, t_c + c\, t_c^2} \right] \tag{4}$$

where

$$b = \frac{B - A}{B\, A} \sigma_0^2 \tag{5}$$

$$c = 0.217 - \frac{\sigma_0^2}{B\, A} \tag{6}$$

The main disadvantage of the above models (Eqs. (1) and (4)) is that the parabolic growth of the variance can lead to non-null values of the baseline for times far from the peak maximum. Thus, for example, for the PVMG model:

$$h(\infty) = H_0 \ \exp\left[-\frac{1}{2c} \right] \tag{7}$$

With the purpose of avoiding the function growth outside the peak region, a modified Gaussian model formed by a combination of a parabole with a Lorentzian (PLMG model) was further proposed [22]:

$$h(t_c) = H_0 \ \exp\left(-\frac{1}{2} \frac{1 + w\, t_c + z\, t_c^2}{\sigma_0^2 + b\, t_c + c\, t_c^2} \ t_c^2 \right) \tag{8}$$

which has demonstrated the best performance in the fitting of isolated asymmetric peaks [7]:

All modified Gaussian models require accurate enough estimations of the height and time at the peak maximum to assure convergence in the fitting process [22,23]. In order to improve the applicability of the model, negative values of the variances should be avoided. It should be also noted that the high flexibility of the PLMG model could give rise to peaks exhibiting more than one maximum.

A large number of common drugs have at least one chiral centre, and the enantiomers may have different pharmacological and toxicological properties. Therefore, the quantification of the enantiomeric fraction, in both commercial and environmental samples, is of great importance [24–26]. The application of all the above models to the description of the peaks obtained in a chiral separation is an important challenge, since the peak characteristics for the enantiomers can be very different and their overlapping degree variable.

The aim of this work is to study the performance of the LMG, PVMG and PLMG models to get reliable fittings of experimental chromatographic peak profiles for several drugs with chiral activity, eluted with enantioselective columns. A main objective is to study how these models succeed in the deconvolution of the chromatographic peaks, when the enantiomers are not completely resolved. This is relevant for a proper evaluation of the enantiomeric fraction of the chiral samples. In this case, the application of multivariate techniques, such as multivariate curve resolution-alternating least squares is not helpful, since both enantiomers in a racemic mixture have the same UV and MS spectra.

2. Experimental

2.1. Reagents

The mobile phases were prepared with acetonitrile (ACN) or methanol (MeOH) (HPLC/LC-MS grade) from VWR International Eurolab (Barcelona, Spain). The following reagents were added to adjust the pH of the mobile phases: ammonium acetate, ammonium

bicarbonate, hydrochloric acid (37%), sodium hydroxide from Scharlau (Barcelona), and ammonium formate from Acros Organics (Geel, Belgium). In some separations, formic acid (98%) and sodium perchlorate monohydrate from Scharlau, and potassium perchlorate, diethylamine (DEA) and potassium hexafluorophosphate from Acros Organics, were used as additives. All reagents were of analytical grade.

Ultra Clear TWF (tap water feed) UV deionised water (SG Water, Barsbüttel, Germany) was used to prepare the following aqueous solutions: 10 mM ammonium acetate and 20 mM ammonium bicarbonate buffers at pH 8.0 adjusted with 2.5 M sodium hydroxide, 10 mM ammonium formate at pH 3.0 adjusted with 1 M hydrochloric acid, 100 mM potassium hexafluorophosphate, and 500 mM sodium or potassium perchlorate. All these solutions were mixed with the organic modifier to obtain the working concentration of the mobile phase. Ternary mixtures of the above aqueous solutions, pure organic modifier and 0.1% (v/v) formic acid or DEA were also used as mobile phases.

The following drugs were analysed: bupivacaine (BUP) from Cayman Chemical Company (Michigan, USA), flurbiprofen (FLU), ibuprofen (IBU), ketoprofen (KET), and trimeprazine (TRI) from Sigma-Aldrich (St. Louis, MO, USA), metoprolol (MET) from Alfa Aesar (Thermo Fisher Scientific, Karlsruhe, Germany), and omeprazole (OME) and propranolol (PRO) from Acros Organics. All the standards used in the experiments were racemates ((±)-analyte). Stock standard solutions of 1000 mg L^{-1} for all the compounds in the study were prepared by dissolving the adequate amount of each drug in MeOH. From these, 100 mg L^{-1} solutions were prepared by dilution in MeOH.

Prior to injection into the chromatographic system, samples were filtered through disposable 0.22 µm polyethersulphone syringe filters (Frisenette, Knebel, Denmark). Mobile phase solutions were vacuum-filtered through 0.22 µm Nylon membranes (Micron Separations,

Westboro, MA, USA), and degassed in an Elmasonic S60 ultrasonic bath (Elma, Singen, Germany) prior to use.

2.2. Apparatus and columns

Two chromatographic systems were used. The first system was an Agilent Technologies 1100 chromatograph (Palo Alto, CA, USA), equipped with a binary pump, a UV-visible diode array detector, a mass spectrometer provided with a source of atmospheric pressure chemical electrospray/ionisation (ESI/APCI) and simple quadrupole, a column thermostat and an autosampler. Data acquisition and processing were performed by means of the LC/MSD ChemStation software (B.04.02 SP1 [208], ©Agilent Technologies 2001–2010).

The second system was an Agilent Technologies 1100 chromatograph, equipped with a quaternary pump, a UV-visible variable wavelength detector, a column thermostat and an autosampler. In this case, data acquisition and processing were performed by means of the Chemstation software (A.09.03 [1417], ©Agilent Technologies 1990–2002).

For the separation of the enantiomers of the drugs under study, several polysaccharide-based chiral stationary phases were used: Lux Cellulose-1 (cellulose tris(3,5-dimethylphenylcarbamate), 3 µm, 150 × 4.6 mm i.d. from Phenomenex, Torrance, CA, USA; Lux Cellulose-2 (cellulose tris(3-chloro-4-methylphenylcarbamate), 3 µm, 150 × 2.0 mm i.d. from Phenomenex; Lux Cellulose-3 (cellulose tris(4-methylbenzoate), 3 µm, 150 × 4.6 mm i.d. from Phenomenex; Chiralart Cellulose-SC (cellulose tris(3,5- dichlorophenylcarbamate), 3 µm, 150 × 4.6 mm i.d. from YMC Separation Technology from Tokyo, Japan; and Lux Amylose-2 (amylose tris(5-chloro-2-methylphenylcarbamate), 3 µm, 150 × 2.0 mm i.d. from Phenomenex.

The mobile phase flow rate was 1.0 mL min^{-1} in all cases, except when using the Lux Cellulose-2 and Lux Amylose-2 columns, for which the flow rate was 0.5 mL min^{-1}. In most cases, column temperature was set at 25 ºC (see Table 1 for more information). The injection

volume was 2 μL, and UV detection of all drugs was performed at 220 nm, except for trimeprazine for which the absorbance was measured at 254 nm. The *m*/*z* ratio was 289 for bupivacaine and 346 for omeprazole, which corresponds to the $[M+H]^+$ ion.

A MicropH 2000 pHmeter (Crison Instruments, Barcelona) was employed to adjust the pH of the buffer solutions.

3. Theory

3.1. Approaches for the prediction of peak profiles

New approaches were developed to eliminate the problem of the prediction of abnormal tails when the chromatographic peaks of enantiomers are fitted using the modified Gaussian models. The approaches are based on the three models described in Section 1 (LMG, PVMG and PLMG). To allow tails tending to zero asymptotically, two modifications of these models are proposed: (i) Replacing the model with exponential tails at 10% peak height outside the peak region, and (ii) applying a sigmoidal function so that the variance tends to a constant value when the time is far from the maximum, giving rise to Gaussian tails. All models should include the half-width values (*A* and *B*) as parameters [1]. In this work, parameters *A* and *B* were measured at 10% peak height to facilitate the application of the proposed modifications.

The different modifications are described below in detail. In all approaches, the variance is forced to be always positive and above 1×10^{-6}.

3.1.1. Approach I: LMG model

Approach Ia: The LMG model is applied directly using Eqs. (1) to (3), without any modification. In this model, the fitted parameters are t_R, H_0, *A* and *B*.

Approach Ib: The peak tails are exponentially modified. The LMG model is substituted by an exponential function below height *p*.

$$p = \frac{h_p}{H_0} \tag{9}$$

To avoid discontinuities, the slope of the LMG function and the slope of the added exponential must be the same at the point of substitution. To simplify, we will write the modified Gaussian function using the general expression:

$$h = H_0 e^{-\frac{1}{2}q^2} \tag{10}$$

whose derivative is:

$$\frac{dh}{dt} = \frac{dh}{dq}\frac{dq}{dt} = -qh\frac{dq}{dt} \tag{11}$$

Meanwhile, the equation of the added exponential is:

$$h = H_0 e^{-kt} \tag{12}$$

and its derivative:

$$\frac{dh}{dt} = -kH_0 e^{-kt} = -kh \tag{13}$$

Matching Eqs. (11) and (13) at point *p*, the decay constant of the exponential function is obtained:

$$k = q\frac{dq}{dt} \tag{14}$$

For the LMG model:

$$q^2 = \frac{t^2}{(\sigma_0 + mt)^2} \tag{15}$$

whose derivative is:

$$q\frac{dq}{dt} = \frac{t\sigma_0}{(\sigma_0 + mt)^3} = \frac{\sigma_0 q^3}{t^2} \tag{16}$$

By substituting Eq. (16) in Eq. (14), the decay constant of the exponential function is as follows:

$$k = \frac{\sigma_0}{t^2} q^3 \tag{17}$$

We will arbitrarily take p = 0.1 (10% peak height) as the point of change of the LMG function to an exponential. Therefore, from Eq. (10):

$$q = \sqrt{-2 \times \ln\left(\frac{h_{0.1}}{H_0}\right)} = \sqrt{-2 \times \ln(0.1)} = 2.146 \tag{18}$$

In this point (p = 0.1), for the left side of the peak ($t_{0.1}$ = -A), the exponential function will be:

$$h = p\, H_0\, e^{k(A+t_c)} = 0.1 \times H_0 \times e^{9.883 \times \frac{\sigma_0}{A^2}(A+t_c)} \tag{19}$$

and for the right side ($t_{0.1}$ = B):

$$h = p\, H_0\, e^{k(B-t_c)} = 0.1 \times H_0 \times e^{9.883 \times \frac{\sigma_0}{B^2}(B-t_c)} \tag{20}$$

Approach Ic: The temporal variable of the variance is replaced by a sigmoidal function:

$$x = \frac{t_c}{1 + 0.2\,\frac{|t_c|}{A+B}} \tag{21}$$

the final peak model being:

$$h(t) = H_0 \exp\left[-\frac{1}{2}\left(\frac{t_c}{\sigma_0 + m\,x}\right)^2\right] \tag{22}$$

This assures that when $t_c = \infty$, then $x = (A+B)/0.2$ and $h = 0$.

3.1.2. Approach II: PVMG model

Approach IIa: The PVMG model is first applied without modifications, using Eqs. (4) to (6). The fitted parameters are t_R, H_0, σ_0, A and B.

Approach IIb: The exponential substitution is applied to the PVMG model, as in Approach Ib. In this case:

$$q^2 = \frac{t^2}{\sigma_0^2 + bt + ct^2} \tag{23}$$

whose derivative is:

$$q\frac{dq}{dt} = \frac{1}{2}\frac{2t \times (\sigma_0^2 + bt + ct^2) - t^2(b + 2ct)}{(\sigma_0^2 + bt + ct^2)^2} = \frac{1}{2}\frac{2\sigma_0^2 t + bt^2}{t^4}q^4 \tag{24}$$

Matching Eqs. (14) and (24), and working out:

$$k = \frac{1}{2}\frac{\left|2\sigma_0^2 t + bt^2\right|}{t^4}q^4 \tag{25}$$

where the absolute value has been taken to ensure a positive value of the decay exponential constant (k). Finally, considering that at 10% peak height, $p = 0.1$ and $q = 2.146$, the equation for the exponential left tail will be:

$$h = 0.1 \times H_0 \times e^{10.604 \times \frac{\left|bA^2 - 2\sigma_0^2 A\right|}{A^4}(A + t_c)} \tag{26}$$

and for the right tail:

$$h = 0.1 \times H_0 \times e^{10.604 \times \frac{\left|bB^2 + 2\sigma_0^2 B\right|}{B^4}(B - t_c)} \tag{27}$$

Approach IIc: As in Approach Ic, the temporal variable of the variance is replaced by Eq. (21). The final peak model will be:

$$h(t) = H_0 \exp\left[-\frac{1}{2}\frac{t_c^2}{\sigma_0^2 + bx + cx^2}\right] \tag{28}$$

3.1.3. Approach III: PLMG model

Approach IIIa: The PLMG approach is applied according to Eq. (8). However, the same as in Approaches I and II, in order to be able to implement Approaches IIIb and IIIc, the half-width values should be known. For this purpose, first of all, parameters *b* and *c* will be related to the half-widths as shown below.

From Eq. (8):

$$q^2 = \frac{1 + wt + zt^2}{\sigma_0^2 + bt + ct^2} t^2 \tag{29}$$

At 10% peak height, $q^2 = 4.605$, and by substituting the time for each half-width ($t_{0.1} = -A$ and $t_{0.1} = B$), we can get two equations from Eq. (29):

$$\sigma_0^2 - bA + cA^2 = 0.217(A^2 - wA^3 + zA^4) \tag{30}$$

$$\sigma_0^2 + bB + cB^2 = 0.217(B^2 + wB^3 + zB^4) \tag{31}$$

Dividing each equation by the corresponding half-width, and adding Eqs. (30) and (31) to eliminate *b*:

$$\sigma_0^2 \frac{A+B}{AB} + c(A+B) = 0.217 \times \left[A + B + w(B^2 - A^2) + z(B^3 + A^3)\right] \tag{32}$$

from which:

$$c = 0.217 \times \left[1 + w(B-A) + z(B^2 + A^2 - AB)\right] - \frac{\sigma_0^2}{AB} \tag{33}$$

Dividing Eqs. (30) and (31) by the corresponding squared half-width, and subtracting to eliminate parameter *c*, the following is obtained:

$$-\sigma_0^2 \frac{B^2 - A^2}{A^2 B^2} + b\frac{A+B}{AB} = 0.217 \times \left[w(A+B) + z(B^2 - A^2)\right] \tag{34}$$

which yields to:

$$b = \sigma_0^2 \frac{B-A}{AB} + 0.217 \times \left[w + z(B-A)\right] AB \tag{35}$$

Eqs. (8), (33) and (35) constitute Approach IIIa, where the model parameters to be fitted are t_R, H_0, σ_0, A, B, w and z. To avoid negative or close to zero values for the peak variance, the parabolas in the numerator and denominator of Eq. (8) will be limited to positive values above 1×10^{-6}.

Approach IIIb: In this approach, exponential tails are added to Approach IIIa. The equations that describe the exponential tails are obtained by deriving Eq. (29):

$$q\frac{dq}{dt} = \frac{1}{2}\frac{2t + 3wt^2 + 4zt^3 - q^2(b + 2ct)}{\sigma_0^2 + bt + ct^2} \tag{36}$$

The decay constant is obtained by matching Eqs. (14) and (36). The left tail is described by:

$$h = 0.1\times H_0 \times \exp\left[\frac{1}{2}\times\left|\frac{-2A + 3wA^2 - 4zA^3 - 4.605\times(b - 2cA)}{\sigma_0^2 - bA + cA^2}\right|\times(A + t_c)\right] \tag{37}$$

and the right tail by:

$$h = 0.1\times H_0 \times \exp\left[\frac{1}{2}\times\left|\frac{2B + 3wB^2 + 4zB^3 - 4.605\times(b + 2cB)}{\sigma_0^2 + bB + cB^2}\right|\times(B - t_c)\right] \tag{38}$$

***Approach IIIc*:** As in Approach Ic, the temporal variable in the variance is replaced by Eq. (21). The final peak model is:

$$h(t) = H_0\ \exp\left[-\frac{1}{2}\frac{1 + wx + zx^2}{\sigma_0^2 + bx + cx^2}t_c^2\right] \tag{39}$$

4. Results and discussion

4.1. Characteristics of the experimental peaks

In order to check the performance of the proposed approaches, 20 chromatograms obtained in the analysis of chiral drugs were selected from our data base. The experimental peaks were obtained using different chromatographic modes, columns, mobile phase compositions, temperatures, and pH values, as indicated in Table 1. The peaks correspond to eight drugs with different pharmacological uses: bupivacaine (BUP1), flurbiprofen (FLU1 and FLU2), ibuprofen (IBU1, IBU2 and IBU3), ketoprofen (KET1, KET2, KET3 and KET4), metoprolol (MET1 and MET2), omeprazole (OME1 and OME2), propranolol (PRO1, PRO2 and PRO3), and trimeprazine (TRI1, TRI2 and TRI3). Peaks of the enantiomeric compounds with diverse efficiency and skewness values were selected. The peaks for BUP1 and OME2 were monitored using both mass spectrometry and UV detection, which decreased the noise in the chromatograms. All other peaks were monitored using UV detection exclusively.

Table 1 gives information about:

(i) The retention times of the peaks for the two enantiomers of each drug.

(ii) The resolution of the two peaks, calculated according to:

$$R_S = \frac{t_{R_2} - t_{R_1}}{B_1 + A_2} \tag{40}$$

where t_{R2} and t_{R1} are the retention times for the peaks at longer (peak 2) and shorter (peak 1) retention, and A_2 and B_1 are the corresponding left and right half-widths measured at 10% peak height.

(iii) The peak asymmetry for each peak (B/A, where peaks are tailing for $B/A > 1$ and fronting for $B/A < 1$).

(iv) The enantiomeric fraction (EF%), measured as the percentage area of peak 1 [25], calculated as:

$$EF\% = \frac{\text{area}_1}{\text{area}_1 + \text{area}_2} \times 100 \quad (41)$$

where area$_1$ and area$_2$ correspond to peaks 1 and 2, respectively.

The ranges of the above parameters for the peaks chosen for this work were: 1.7–27.5 min for the retention times, 0.49–3.34 for peak resolution, 0.84–3.13 for peak asymmetry, and 29.8–61.2% for the enantiomeric fraction of the first eluted enantiomer. All values were measured from the individual peaks obtained from the fittings according to Approach IIIb. In Fig. S1 in the Supplementary material, the chromatograms for the probe compounds are shown.

4.2. Fitting performance

The approaches described in Section 3 (LMG, PVMG and PLMG, without and with restrictions) were applied to fit the peaks in the chromatograms for the enantiomeric compounds (Fig. 1). It should be noted that the peaks of the two enantiomers were fitted simultaneously:

$$\hat{h}_i = h_1(t_i) + h_2(t_i) \quad (42)$$

h_1 and h_2 being the height of peaks 1 and 2, respectively at time t_i, and $\hat{h}_i$ the predicted signal at this time, which is the summation of the contribution of both peaks. The fitting is performed using least squares regression by optimising the peak parameters of each peak. For instance, for the LMG model, eight parameters were optimised: t_{R1}, H_{01}, A_1 and B_1 for peak 1, and t_{R2}, H_{02}, A_2 and B_2 for peak 2.

When the peaks of the enantiomers are overlapped, they can only be resolved by making the simultaneous fitting of the functions of both enantiomers (Eq. (42)). Of course, there is no advantage in doing so when the peaks are perfectly spaced, more than saving time. However, this allows obtaining a measurement of the quality of the fittings common to all peak pairs, independently of their resolution.

The fitting performance was assessed by calculating the mean relative errors according to:

$$\varepsilon_r(\%) = \frac{\sum_{i=1}^{N} \left| h_i - \hat{h}_i \right|}{\sum_{i=1}^{N} \left| h_i \right|} \times 100 \quad (43)$$

where h_i and $\hat{h}_i$ are the experimental and predicted heights for each point in the chromatogram (at different times), and N is the number of experimental points.

Fig. 1 illustrates the performance of the different approaches based on the LMG, PVMG and PLMG models, without any restriction (Approaches Ia, IIa and IIIa), and after applying the restrictions described in Section 3 (Approaches Ib, IIb and IIIb, and Approaches Ic, IIc and IIIc). The excellent fitting achieved by applying the PLMG model for the enantiomeric peaks in KET4, IBU3 and MET2 (which show a large diversity in peak tailing, from B/A = 0.84 to 2.81), is remarkable. Similar behaviour was observed for the other peaks studied in this work.

The errors obtained in the fitting of the chromatograms of the chiral drugs studied in this work, according to each approach for the LMG, PVMG and PLMG models (without and with restrictions), are indicated in Table 2. As observed, the errors for the PLMG model were below 1% for most compounds. Approaches IIIa, IIIb and IIIc showed a mean relative error of 0.69%, 0.80% and 0.71%, considering the 20 studied peak-pairs, respectively. These errors were significantly smaller than the errors yielded by fitting the peaks according to the LMG and PVMG models. The behaviour was acceptable when the PVMG model was applied, with mean relative errors of 1.87%, 1.52% and 1.58% for Approaches IIa, IIb and IIc, respectively. The worst results were obtained with the LMG model, which is the simplest among those studied, with mean relative errors in the fitting of the chromatograms of 4.34%, 4.12% and 4.54% for Approaches Ia, Ib and Ic, respectively. Note also that for all studied cases, there was a notable improvement in the performance of the fittings, when the PVMG model was used instead of the LMG model.

The slight improvement in the fittings obtained with the PVMG model when exponential tails were added (Approach IIb), or the growth of the independent variable was limited (Approach IIc), should be also highlighted. Meanwhile, Approach b causes also a slight improvement for the LMG model, and a slight worsening for the PVMG model. For Approach c, the results of the LMG and PLMG models are worse.

However, in spite of the apparent better performance of the PLMG model, it should be indicated that this model is too flexible and may incur in overfitting in those regions where the peaks of the enantiomers are overlapped. This is observed in Fig. 2, where the experimental chromatogram for KET4 is overlaid with the individual peaks obtained through the fitting with the PLMG model, without restrictions (IIIa) and using the two proposed restrictions (IIIb and IIIc). Approaches IIIa and IIIc gave rise to anomalous peak shapes in regions where the peaks of the enantiomers were overlapped. Approach IIIb solved this anomaly by forcing the peak tails to follow an exponential decay.

The lack of reliability could persist if the overlap occurs above 10% peak height, a situation where the model may even predict peaks showing two maximums. It should be noted that the flexibility of the PLMG model is excessive for overlapped compounds, and the reliability of the fitted parameters A and B is low. Therefore, the PLMG model cannot be used for the deconvolution of overlapped peaks, without being constrained to a proper peak shape. However, it is useful for the fitting of perfectly defined individual peaks, to get a peak function describing accurately the experimental data. Outside the peak region, the function value should be set to zero to avoid problems with the baseline. This function can be thus used without restrictions to obtain peak parameters or reduce noise problems.

From the results obtained, it can be seen that Approaches IIb and IIc provide very similar results, Approach IIc being easier to apply. This approach will be used in the next section to study the deconvolution of overlapped peaks, together with the PLMG model (Approach IIIa).

4.3. Peak deconvolution

As commented, the PLMG model is the best to obtain the characteristics of an isolated peak [1]. It is very flexible and can fit very satisfactorily perfectly profiled peaks in a wide range of peak shapes, inside specific time intervals. However, it could be not reliable enough to deconvolve overlapped peaks (as could be the case of some enantiomeric compounds), or to make predictions at times outside the fitted range.

Once the baseline effects have been removed, for example with the exponential decay of Approaches Ib, IIb and IIIb, the difference in the reliability of the deconvolution depends on the capability of the model to extract the correct peak shape of each enantiomer. This is possible if the experimental errors are random Gaussian and the peak shape is maintained constant identical to the individual peaks [22]. However, in the case of enantiomeric mixtures, the individual peak information is not easy to obtain.

We will first compare the stability of the two models that showed the best performance in the fitting of chromatographic peaks (PLMG and PVMG). In this study, both models were used to fit the overlapped peaks of KET4, for samples injected at different concentrations (2, 5, 10, 30 and 100 mg L^{-1}). In Fig. 3a, the experimental chromatograms are shown. With the information obtained in the deconvolution, the area of each peak was measured in order to calculate the enantiomeric fraction.

The deconvolution of the peaks for the chromatograms obtained at different concentrations should show the same enantiomeric fraction, since the peak shape is not expected to change. Fig. 3b compares the results obtained using Approaches IIc (PVMG model) and IIIa (PLMG model) to fit the chromatograms. The plot shows the large stability in the values of enantiomeric fraction when the PVMG model was applied (i.e., the change in the calculated fraction for chromatograms at different concentrations was rather small). The variation in the calculated enantiomeric fraction for the same chromatograms was appreciably larger for the PLMG model.

This can be explained by the large uncertainties in the peak shapes with this model, and consequently, the larger errors in the deconvolution of the overlapped peaks. In contrast, both PVMG and PLMG models yielded good fitting of the chromatograms, resulting in similar total areas with a mean relative error of 0.3%. Therefore, it is evident that the PLMG model makes an incorrect distribution of the total area between the two individual peaks.

Finally, we considered interesting to study the behaviour of the PVMG model when it is applied to chromatograms containing overlapped peaks. Fig. 4 shows the chromatograms obtained for KET2 (R_S = 0.73) and KET3 (R_S = 1.02), together with the predicted individual peaks obtained with the PVMG model. It should be noted that the evaluated enantiomeric fractions, using Approach IIc is 38.3% for KET2 and 46.4% for KET3, the latter being better resolved. As observed, peak 2 increased its width in KET2, incorporating part of the area of peak 1. This may be due to the variation in the width of one peak along the fitting iterations to compensate the decrease in the width of the other, to get a better fit. On the other hand, significant changes could occur in the individual peaks when eluted overlapped along the chiral column. For KET3, this effect was not observed owing to the better resolution.

This shows that the PVMG model does not yield reliable deconvolutions when all peak parameters are fitted and the resolution is poor (below R_S = 1.0 for the data set studied in this work). Further study should be done to overcome this problem, limiting the peak shape variation. For example, if we keep fixed the parameters of the PLMG model, except for t_R, H_0, A and B, we would have a model with four parameters with a fairly stable peak shape as in Approach I. This can also be observed in Figs. S2 to S5 depicted in the Supplementary material. When the resolution is sufficient (1.19 and 1.38 for Figs. S2 and S5, respectively), all the approaches give similar EF values. However, in Figs. S3 and S4 with resolutions of 0.49 and 0.62, respectively, the variability is larger. Again, Approaches I (where only the shape parameters A and B are fitted) show greater similarity, and Approaches III (where five shape

parameters are fitted, give rise to larger variability. Obviously, it is necessary to find ways to limit the peak shape with Approach III, without losing goodness of fit.

5. Conclusions

In this work, modifications of some modified Gaussian models are proposed and evaluated. We have seen that in order to obtain adequate results for overlapping peaks, it is necessary to restrict the models to set the peak shape in a proper way. The family of functions of modified Gaussian peaks allow the fitting of chromatographic peaks with a wide range of properties, from symmetric to highly asymmetric. The deconvolution of the components of a mixture of enantiomeric compounds insufficiently resolved with a chiral column is also possible. In this work, different approaches based on the LMG, PVMG and PLMG models were applied to fit a variety of peaks of enantiomeric compounds of eight drugs, separated with several chiral columns. The simplest function (LMG) provided satisfactory fittings only for sufficiently symmetric peaks. The PVMG function gave rise to adequate fittings for most peaks, being the model of choice for compounds eluted in the optimal conditions, when the peaks show conventional profiles. However, the best fittings were obtained with the PLMG model, although in some situations overfitting problems or unsatisfactory tails may arise. Even eliminating the tails by adding an exponential function, the overfitting did not disappear, which may be problematic for the deconvolution of overlapped peaks.

The application of exponential tails (Approaches Ib, IIb and IIIb) is an appropriate option to eliminate problems, but does not significantly improve the fittings. On the other hand, the LMG and PVMG models using a sigmoidal function to limit the growth of the temporal variable in the variance (Approaches Ic and IIc) provide acceptable results and are easier to apply.

We have observed that in order to obtain adequate results for overlapping peaks, it is necessary to restrict the models to set the peak shape. When the models are used without

restrictions, they tend to overfitting, and transfer an area fraction from one peak to another if this yields a better fit. A more extensive study of the peak shape variations is necessary for poorly resolved enantiomers. Also, a procedure should be developed to keep the enantiomeric fraction constant in the deconvolution procedure, which will be the purpose of future work. The proposed models should be still tested when the enantiomeric fraction is far from racemic.

Acknowledgements

This work was supported by Projects CTQ2015-70904-R (Ministerio de Economía y Competitividad, Spain), CTQ2016-75644-P (Ministerio de Economía, Industria y Competitividad, Spain), and PID2019-106708GB-I00 (Ministerio de Ciencia, Innovación y Universidades), and FEDER funds, and PROMETEO/2016/128 (Direcció General d'Universitat, Investigació i Ciència, Generalitat Valenciana, Spain). Mireia Pérez- acknowledges Generalitat Valenciana and European Social Fund for the contribution to the pre-doctoral contract ACIF/2019/158.

References

[1] M.C. García-Alvarez-Coque, J.R. Torres-Lapasió, J.J. Baeza-Baeza, Models and objective functions for the optimisation of selectivity in reversed-phase liquid chromatography, Anal. Chim. Acta 579 (2006) 125–145.

[2] J.R. Torres-Lapasió, M.C. García-Alvarez-Coque, Levels in the interpretive optimisation of selectivity in high-performance liquid chromatography: a magical mystery tour, J. Chromatogr. A 1120 (2006) 308–321.

[3] R. Cela, E.Y. Ordoñez, J.B. Quintana, R. Rodil, Chemometric-assisted method development in RPLC, J. Chromatogr. A 1287 (2013) 2–22.

[4] A. Periat, D. Guillarme, J.L. Veuthey, J. Boccard, S. Moco, D. Barron, A. Grand-Guillaume-Perrenoud, Optimized selection of liquid chromatography conditions for wide range analysis of natural compounds, J. Chromatogr. A 1504 (2017) 91–104.

[5] J.A. Navarro-Huerta, E.J. Carrasco-Correa, J.R. Torres-Lapasió, J.M. Herrero-Martínez, M.C. García-Alvarez-Coque, Modelling retention and peak shape of small polar solutes analysed by nano-HPLC using methacrylate-based monolithic columns, Anal. Chim. Acta 1086 (2019) 142–155.

[6] J.J. Baeza Baeza, M.C. García-Alvarez-Coque, Prediction of peak shape as a function of retention in reversed-phase liquid chromatography, J. Chromatogr. A 1022 (2004) 17–24.

[7] J.J. Baeza-Baeza, C. Ortiz-Bolsico, M.C. García-Alvarez-Coque, New approaches based on modified Gaussian models for the prediction of chromatographic peaks, Anal. Chim. Acta 758 (2013) 36–44.

[8] M.J. Ruiz-Angel, S. Carda-Broch, M.C. García-Alvarez-Coque, Peak half-width plots to study the effect of organic solvents on the peak performance of basic drugs in micellar liquid chromatography, J. Chromatogr. A 1217 (2010) 1786–1798.

[9] J.J. Baeza-Baeza, M.J. Ruiz-Angel, M.C. García-Alvarez-Coque, S. Carda-Broch, Half-width plots, a simple tool to predict peak shape, reveal column kinetics and characterise chromatographic columns in liquid chromatography: state of the art and new results, J. Chromatogr. A 1314 (2013) 142–153.

[10] E. Grushka, Characterization of exponentially modified Gaussian peaks in chromatography, Anal. Chem. 44 (1972) 1733–1738.

[11] J.P. Foley, J.G. Dorsey, Equations for calculation of chromatographic figures of merit for ideal and skewed peaks, Anal. Chem. 55 (1983) 730–737.

[12] J.P. Foley, Equations for chromatographic peak modeling and calculation of peak area, Anal. Chem. 59 (1987) 1984–1987.

[13] J.R. Torres-Lapasió, J.J. Baeza-Baeza, M.C. García-Alvarez-Coque, A model for the description, simulation, and deconvolution of skewed chromatographic peaks, Anal. Chem. 69 (1997) 3822–3831.

[14] P. Nikitas, A. Pappa-Louisi, A. Papageorgiou, On the equations describing chromatographic peaks and the problem of the deconvolution of overlapped peaks, J. Chromatogr. A 912 (2001) 13–29.

[15] K. Lan, J.W. Jorgenson, A hybrid of exponential and Gaussian functions as a simple model of asymmetric chromatographic peaks, J. Chromatogr. A 915 (2001) 1–13.

[16] V.B. Di Marco, G.G. Bombi, Mathematical functions for the representation of chromatographic peaks, J. Chromatogr. A 931 (2001) 1–30.

[17] J. Li, Comparison of the capability of peak functions in describing real chromatographic peaks, J. Chromatogr. A 952 (2002) 63–70.

[18] S.V. Romanenko, A.G. Stromberg, T.N. Pushkareva, Modeling of analytical peaks: Peak properties and basic peak functions, Anal. Chim. Acta 580 (2006) 99–106.

[19] J.R. Torres-Lapasió, J.J. Baeza-Baeza, M.C. García-Alvarez-Coque, Modeling of peak shape and asymmetry, in Chemometrics in Chromatography, Vol. 111 de Chromatographic Science Series (edited by L. Komsta, Y. Vander Heyden, J. Sherma), pp. 217–238 (2018), CRC Press, Boca Raton, FL.

[20] J.P. Foley, J.G. Dorsey, Equations for calculation of chromatographic figures of merit for ideal and skewed peaks, Anal. Chem. 55 (1983) 730–737.

[21] J.P. Foley, J.G. Dorsey, A review of the exponentially modified Gaussian (EMG) function: evaluation and subsequent calculation of universal data, J. Chromatogr. Sci. 22 (1984) 40–46.

[22] R.D. Caballero, M.C. García-Alvarez-Coque, J.J. Baeza-Baeza, Parabolic-Lorentzian modified Gaussian model for describing and deconvolving chromatographic peaks, J. Chromatogr. A 954 (2002) 59–76.

[23] P.G. Stevenson, F. Gritti, G. Guiochon, Automated methods for the location of the boundaries of chromatographic peaks, J. Chromatogr. A 1218 (2011) 8255-8263.

[24] Lien Ai Nguyen, Hua He, Chuong Pham-Huy, Chiral drugs: an overview, Int. J. Biomed. Sci. 2 (2006) 85–100.

[25] A.R. Ribeiro, L.H.M.L.M. Santos, A.S. Maia, C. Delerue-Matos, P.M.L. Castro, M.E. Tiritan, Enantiomeric fraction evaluation of pharmaceuticals in environmental matrices by liquid chromatography-tandem mass spectrometry, J. Chromatogr. A 1363 (2014) 226–235.

[26] L. Escuder-Gilabert, Y. Martín-Biosca, M. Pérez-Baeza, S. Sagrado, M.J. Medina-Hernández, Direct chromatographic study of the enantioselective biodegradation of ibuprofen and ketoprofen by an activated sludge, J. Chromatogr. A 1568 (2018) 140–148.

FIGURE CAPTIONS

Figure 1. Experimental (○) and predicted chromatograms (dashed line - - -) for ketoprofen (KET4), ibuprofen (IBU3), and metoprolol (MET2). Predicted peaks (black line, —) obtained by applying the studied approaches are overlaid. Only a few experimental points are shown.

Figure 2. Performance of the PLMG model in the fitting of the experimental signal of ketoprofen (KET4) (dashed line - - -). Predicted peaks (black line, —) obtained by applying Approaches IIIa, IIIb and IIIc are overlaid.

Figure 3. (a) Deconvolution of KET4 data eluted at several concentrations: 2, 5, 10, 30 and 100 mg L^{-1} (from bottom to top). (b) Enantiomeric fraction obtained with Approach IIIa (model PLMG) (●) and Approach IIc (model PVMG) (○).

Figure 4. Performance of the PVMG model in the fitting of the experimental signals of ketoprofen (KET2 and KET3) (dashed line - - -). Predicted peaks (black line, —) obtained by applying Approach IIc are overlaid.